Crime and Incident Watch for Smart Cities: A Sensor-Based Approach

Francis N. Nwebonyi(✉), Xiaoyu Du, and Pavel Gladyshev

Digital Forensics Investigation Research Laboratory, School of Computer Science, University College Dublin, Dublin, Ireland
{francis.nwebonyi,xiaoyu.du,pavel.gladyshev}@ucd.ie
http://dfire.ucd.ie/

Abstract. Beyond the many advantages which Smart City brings, the issue of security and privacy remain very important concerns. As things get more interconnected, cyber orchestrated crimes such as cyberterrorism may become more prevalent. For example, an autonomous vehicle or drone may be used to commit acts of terrorism, such as driving or flying into a crowd. Responses to such incidents need to be swift and likewise smart. Law enforcement needs to be equipped with the necessary tools to respond speedily and even automatically, equally taking advantage of the Internet of Things (IoT). We propose a Sensor-Based Crime and Incident Watch for Smart Cities (SBCI-Watch), leveraging Human Activity Recognition (HAR). SBCI-Watch may be used to automatically detect and possibly report occurrences of public disturbances (likely caused by acts of terrorism or similar crimes) to foster swift response. Slightly similar reports exist, but they are focused on different topics, such as helping law enforcement to pick suitable officers to respond to crime scenes based on proximity, but without attempting to automatically detect crimes. We are using the sensor-based approach; it is less privacy intrusive compared to the vision-based method, which is dominant in the public surveillance area but also more privacy intrusive and more expensive. To the best of our knowledge, our work is the first to address the problem in this fashion, and the results illustrate the viability of the SBCI-Watch.

Keywords: Smart City · Human Activity Recognition (HAR) · Incident watch · Crime watch · Privacy

1 Introduction

We can tell the behaviour of an entity or object by observing its movement pattern. By looking at the variation of the speed of cars, for instance, we can get an idea of the traffic pattern. Significantly high variation may mean that the vehicle slows down frequently after picking up some speed, and that can give us some information about the traffic situation. A smoother transition and less variation may suggest a smoother traffic condition.

S. Goel et al. (Eds.): ICDF2C 2022, LNICST 508, pp. 222–238, 2023.
https://doi.org/10.1007/978-3-031-36574-4_13

It can also be applied at the pedestrian level to predict people's activities in public spaces by observing their movement patterns. The literature documents these diverse applications, including the application in the area of transportation systems [5], traffic-anomaly-detection; that is, observing when there are traffic obstructions due to large events, protests, etc. [6], predicting or sensing people's activities such as shopping, working, in a meeting or at home [7], and detecting abnormal activities in social events [8].

The obvious and popular method for observing activities that are indicative of criminal acts such as acts of terrorism and similar emergencies in the public space is video surveillance, using CCTV cameras and similar devices. Although such surveillance devices may be common, it can be very expensive to hire people with the right skill-sets to analyse the videos. This has led to the popularity of real-time analysis methods, to better detect the activities and thus deter and/or respond accordingly. Many video-based algorithms for detecting abnormal behaviours in the crowd have therefore been proposed in recent years to improve the detection of such abnormal behaviours in the crowd; each method attempting to address a particular challenge associated with the process. These challenges range from crowd occlusion to illumination challenges, especially in a dense crowd [4]. There are other issues with the visual method of surveillance as we shall discuss shortly.

In this work, we have adopted a different approach, by extending the non-video-based behaviour observation methods, which have been applied in other areas, towards improving public safety by using it as a kind of a surveillance system, particularly in the context of a smart city, and not with the goal of replacing the video-based surveillance but rather to complement and extend its reach. We shall now briefly discuss the concept of Human Activity Recognition (HAR) which is the umbrella concept for this topic; we will also further highlight the relevance of our approach. The remainder of the paper discusses HAR as part of Sect. 1 and literature review in Sect. 2. The Proposed Approach and Methods are enumerated in Sect. 3, before experiment and results in Sect. 3.2, and conclusion and future work in Sect. 4.

Human Activity Recognition (HAR): This concept has gained popularity due to its application in various aspects of the society, including event analysis, Ambient Assisted Living (AAL), Intelligent Video surveillance systems, and anomaly detection, among others. We can detect anomalies in human activities by identifying non-regular or non-conforming patterns in a given data; the more distinctive the activity is, the more easily detectable it may be. Motion orientation and magnitude are among the common attributes that are often used.

There are two main approaches to HAR; vision-based, and sensor-based [27]. The vision-based approach uses RGB video cameras, as well as depth cameras to capture human actions and skeletal data. On the other hand, the sensor-based approach uses sensors such as body-worn devices, ambient, mobile phone sensors, etc. The vision-based is also popularly categorized as single-person, and multiple-person or crowd, and then applied based on need. For instance, single-person techniques like silhouette have been applied in the area of elderly healthcare to observe falling, fainting, chest pain, headache, etc. Despite being popular, the silhouette has also been reported to have a high cost of implementation, and unsuitable for long-term real-time analyses [1,2].

Multiple-person HAR or simply Crowd Activity Analysis is focused on identifying human activities in the crowd, e.g. riots and acts of terrorism. The first step under this method is usually to identify patterns of group activities, in order to recognize when a given activity becomes abnormal. And the second step is often to represent behaviours of individuals in the crowd using trajectory-based and shape-based techniques, among others.

In general, the vision-based HAR raises very high privacy concerns, way more than the sensor-based methods. However, there is a third method which is also beginning to gain popularity and is considered even less intrusive than sensor-based approach; the WiFi-based method. The bases for the WiFi-based method is that a human activity can generate multi-path reflections of wireless signals from a transmitter (TX) to a receiver (RX), which can be uniquely used to identify such activity. The Receiver Signal Strength (RSS) and the Chanel State Information (CSI) are the two popular WiFi-based HAR approaches. Although the CSI may withstand more environmental variations compared to RSS (which is single-valued), WiFi-based HAR is generally more suitable for well-controlled and fixed environments, e.g. indoors. For this reason, it may not be suitable for the proposed system, which is why we have opted for a sensor-based approach [3].

The real-world surveillance system is dominated by the vision-based HAR. But apart from the earlier mentioned privacy concerns, there are two more key problems with this method; (i) Cost of installation and upgrading of video capturing devices: There are millions of surveillance cameras all across Europe, for example, but they still do not cover every part because it may not be financially efficient to do so. Additionally, upgrading the equipment to keep up with the advancements in technology can also be very expensive. Currently, there are many older surveillance systems that are of very low resolution and are in need of upgrade. (ii) The footage of the crowd in public places is often self-occluding, and as a result, it can be difficult to accurately describe captured individuals and their actions [4]. Sensor-based devices such as the mobile phone may help us to overcome these challenges since individuals own their phones and smart devices, thereby flattening the curve on the cost of purchasing hardware surveillance equipment. The issue of self-occluding is also resolved; since there are no videos involved.

2 Related Work

The traffic system that we currently have is a result of huge research efforts dating back to 1970 and earlier. At that time, traditional means such as surveys, slow motion pictures, trial and error experimentation were used for the research and studies in this area, as demonstrated by Fruin JJ. in his thesis [10]. But modern monitoring systems have since emerged, and are being adopted, including surveillance cameras, Bluetooth, WiFi, and Sensor based techniques.

Yang et al. [11] studied the movement characteristics of pedestrians in staircase, and found that their speed differs considerably under normal and emergency situations, while Wang et al. and others have worked on identifying abnormal behaviours in the crowd [12–14], albeit using the vision-based approach. In this work, however, we shall be focusing on the application of sensor-based HAR in related areas.

Application of Sensor-Based HAR in Museums, and for Crowd Management: Understanding the visitors' dynamics in the Museums can help the management to make better plans. For instance, peak periods and art pieces which attract higher attention can be identified for appropriate planning. Centorrino et al. used Bluetooth beacons for this purpose. Visitors are given the beacons upon entrance, while stationed receivers capture the trajectory of the visitors (can also be called pedestrians) as they move within the museum [15]. Similar techniques are also used in large events to keep track of visitors or recognize returning ones, and monitor passengers' movements in a train station [16,19].

WiFi Access points are also used to gather location-related information in similar environments. For instance, Danalet et al. used this technique to gather data and model the choice of catering locations on a campus. As people login to use the WiFi, anonymized data were gathered and processed [17]. A similar method was used by Gioia et al. to monitor dynamic crowds and to gather information that may be relevant for planning, such as number of people in attendance as well as their spatiotemporal distribution [18].

Organizing large events can be tedious, and as crowds move within a 'confined' space, there may be a risk of a stampede in cases where appropriate plans are not made. To help event planners and crowd managers, Blanke et al. [21] leveraged HAR by using the GPS traces from mobile phones to monitor (or predict) crowd movement in a big event, focusing on schedules and attraction locations within the crowd mobility. Also based on GPS traces and still in the concept of crowd management, Duives et al. [22] proposed a model for forecasting the movement of people in a crowd, leveraging Recurrent Neural Networks. The area of crowd management and mobility has received significant attention, and information from mobile phones and other sensors have been leveraged to address concerns in this area [8,23,24].

Application of Sensor-Based HAR on Traffic Management and City Centre Planning: To observe and investigate the movement and experience of pedestrians in a city centre, Van in his work distributed GPS tracking devices to visitors in a city, at strategic locations such as parking lots. The devices were returned by the volunteers after moving about with it, and the gathered information was further analysed [20]. The goal of this project was to help the city managers to improve the physical condition and experience of visitors in their city centres. It was intended to help them in the areas of beautification and landscaping, among others.

The sensor-based HAR has also been featured considerably in the area of traffic management. The authors of [5,6,25,26] have utilized the trajectory information, gathered while using taxis as sensors, to monitor the traffic behaviours and (when possible) suggest alternative routes to drivers. By recognizing the information that is related to the vehicles as they travel along different routes, researchers are able to predict or suggest safer routes to users or volunteers. It is also possible to fish out a malicious taxi operator who would rather take passengers through an unusually long route to a destination that has a clearly shorter route, perhaps to charge a higher fare.

And for indoor or in-campus navigation, Jackermeier et al. [28] worked on being able to recognize when a pedestrian is performing activities like walking straight, walking through a door by pushing or pulling it, turning right, left, standing still or just looking around. Similarly, Faye et al. [7] used the GPS and accelerator sensors in mobile phones and smartwatches to understand the activities of volunteers, such as when they are at work, shopping, etc. The information provided by the sensor allowed the researchers to determine if the participants are sitting, running, walking or in a vehicle. Such information was then correlated with activities such as shopping, working at the office and similar others.

Public Safety and Policing the Smart City: Joh has postulated that as cities become smarter, they also embed policing within themselves in an increasing way [29]. Smart city features (eg. HAR as we have seen here) can improve service delivery in several areas such as traffic management, event management, parking space, among others. Joh argues that if it can improve services in those areas, then it can also enhance even a more important service which is the public safety and policing. There may be a few laws limiting data collection for use by law enforcement, but Josh has noted that in the case of the United States, for instance, data collection about persons and activities in public spaces is not against the first Amendment. And if the law enforcement retrieves data from third-party companies, then not many restrictions would apply under the current American law, according to Joh. This may suggest that the use of the SBCI-Watch would be in line with existing laws in America and other western countries. Nonetheless, we assume that the responsibility of ensuring the fulfillment of every legal or regulatory requirement for data use will rest upon a trusted third party. Regulatory matters are therefore out of scope here. The data which we have used for experiment was collected under full consent and was anonymized.

Applying sensor-based HAR for crime watch or policing-related goals are emerging and there are no much work in the literature at the moment. A somewhat related work is reported by Welsh and Roy [30]; using sensors that are built into smartphones, they worked on recognizing gunshots, and their model achieved an average accuracy of 86.6% on gunshot classification. But as mentioned earlier, the focus of this and similar reports are different. Our work focuses on utilizing non-vision-based HAR to detect public disturbances or serious crimes such as acts of terrorism, so that appropriate authorities may be promptly notified, leading to prompt response. We utilize common sensors built into smartphones and Apps which people already use, to address this. To the best of our knowledge, we are the first to address this problem using the mentioned approach.

3 The Proposed Approach and Methods

In this section, the concept and overview of the SBCI-Watch are presented, as well as the data collection method and the modeling approach.

3.1 The Concept of SBCI-Watch

As highlighted in Sect. 2, although the smart city may present novel security challenges when rolled out at a bigger and more detailed scale, it also comes equipped with resources that may help to bring about the solution to that problem - mainly the advanced interconnectedness and data pool. The proposed system aims at taking advantage of these resources to better secure the city. We consider it important to focus on utilizing the readily available resources that are already embedded in the city to bring about the needed solution. Because this would make data management, as well as the adoption of the proposed solution, easier. Popular Apps and smart devices are part of the rich resources of smart cities.

Figure 1 illustrates the SBCI-Watch concept. It starts with the gathering of anonymous information from consenting individuals using a mobile App. The information is then passed onto a deep learning model which is able to predict the movement pattern of people in a given location (or venue, eg. Museum) at a given time. The idea is that if a reasonably large number of people within a location or venue (in a smart city) at a given time suddenly start running in an unusual and chaotic way, clearly different from the average known pattern of that particular area or location (see Fig. 3), then perhaps something may be going on and some attention may need to be called to it. An example could be if there is an active shooter or armed rubber at a location and many people suddenly start running for their lives away from the dangerous spot, towards different other directions. Having information about past patterns for that location or venue, and at a similar time/season can help in determining such sudden chaotic scene.

In addition to understanding the movement pattern, other factors, such as analysing whether or not there has been shooting, car accidents, collusion, etc. in the area can provide further verification steps and more confidence in the conclusion which the model may arrive at. Information about firearms has been incorporated into SBCI-Watch as part of the verification step (also see Fig. 5).

Once such information has been gathered, the next step, as also portrayed in Fig. 1, is to pass the data through our deep learning model (which will be residing with a trusted authority) for analysis. The model uses the information which can be retrieved from the data to determine whether there is a suspected occurrence of acts of public disturbance or a related crime, and if that is the

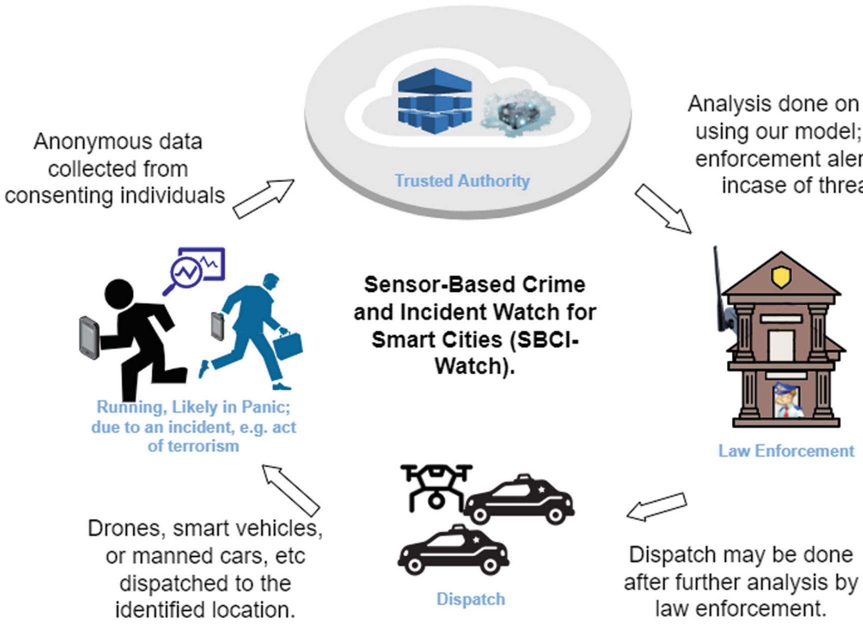

Fig. 1. An Illustration of a Sensor-Based Crime and Incident Watch for Smart Cities (SBCI-Watch)

case, the appropriate authorities (e.g. the law enforcement) can be alerted to further scrutinize the report and dispatch response if it is deemed appropriate. Some of the key steps which we take to reach a tentative conclusion regarding whether there is an incident or not is illustrated in Fig. 2, while the analysis and verification part is further illustrated in Fig. 5 and discussed in Sect. 3.2.

The Long Short-Term Memory (LSTM) has been used to train our model because information about past events needs to be factored in, in order to determine if there has been a public disturbance that is worth being flagged. For instance, while analysing the state of a location or venue at a given time, we need to also take into account what the situation of that area has been prior to the time of the suspected incident. LSTM is known to be very good at this kind of tasks, ensuring that while making analysis at time t, the occurrences at t-1, t-2, etc. are not ignored, because they help to determine unusual patterns.

3.2 Experiment and Results

As earlier mentioned, vision-based HAR is the predominant approach in this area, traditionally involving CCTV cameras and similar other devices, which are more privacy intrusive and expensive to install, maintain and upgrade. But we can also detect public disturbances and crime in public spaces by understanding people's reactions and movement patterns, in addition to other information, which can all be gathered with common sensors that are present in mobile phones

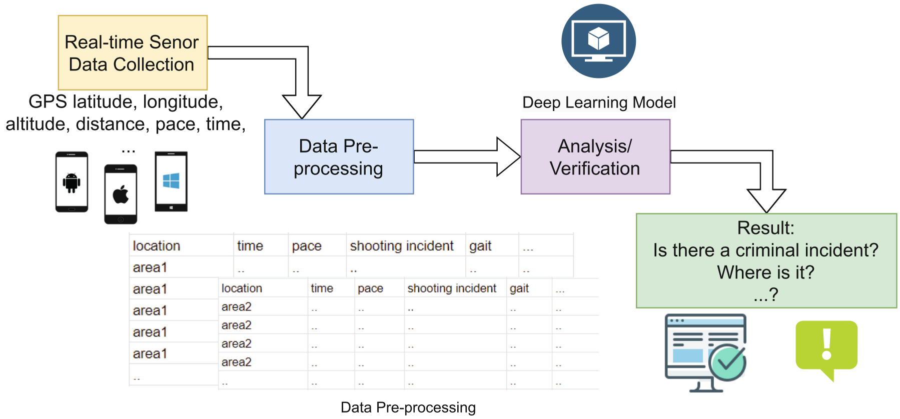

Fig. 2. Data Analysis Steps.

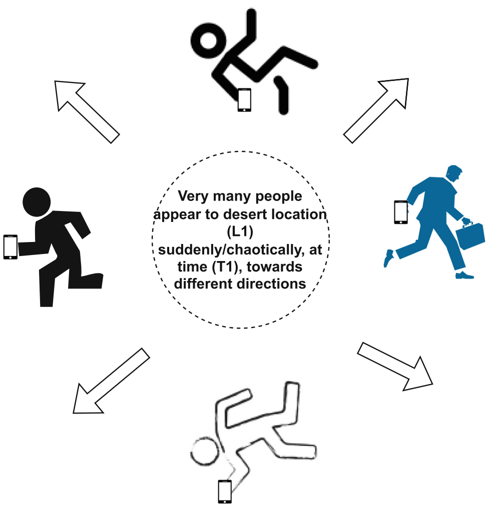

Fig. 3. People running in a chaotic fashion, away from a suspected incident or crime.

and other smart devices; this is the focus of our work. Nonetheless, due to the difference in focus between our work and earlier reports of HAR in the literature, the existing dataset is not suitable for the SBIC-Watch use case, which is why we assembled a fresh dataset. We shall however incorporate features from a firearm dataset [33] into the model as an added verification layer.

Generating the Data Set: In line with our idea of using already available resources, we used an App that is already available at the Play Store and App-store, called GPS Logger. The interface and settings are shown in Fig. 4. Some of its features include background logging (GPS latitude, longitude, altitude, distance, speed) which is more power-friendly compared to Apps that require active recording. The GPS Logger also allows files to be exported in different file formats including CVS which makes it almost ready for instant pre-processing.

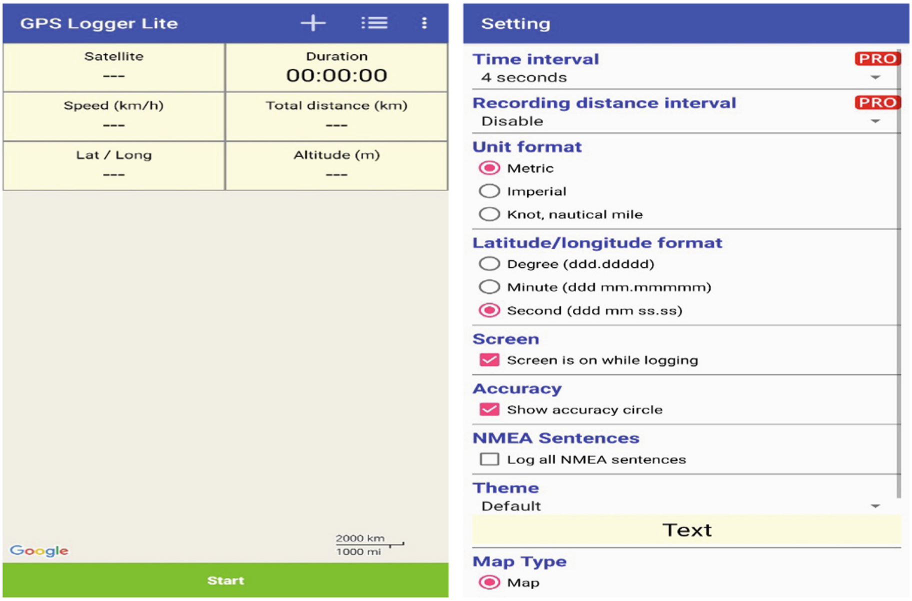

Fig. 4. GPS Logger and settings

A total of seven people participated in the data collection exercise, and it was mainly collected in 2 days involving many sessions. First with 3 volunteers and then with six volunteers on the second day. Two of the volunteers participated on both exercises. Varieties of phone models were used but they were all either Android or iOS; they include iPhone 11 Pro Max, OnePlus 6T, iPhone 7 Plus, Samsung Galaxy A51, Infinix 8, and Samsung Galaxy J5. Participants walked at their normal paces from different directions, heading towards a location. And before they were to reach a common location, they were "startled" with a loud sound as a way of simulating a sudden chaotic situation. On hearing that sound they ran (as they would in a panic situation) in different directions. This way, the area or spot where an attack or public disturbance has occurred maybe predictable.

The experimental process was explained to the participants, including the use of a sound to simulate a situation of panic. The App which was installed on the phones were turned on before each session. The process was repeated many times, and we recorded information about the time, location (latitude/longitude), altitude, pace/speed (km/h), and total distance (km). Data was generally captured every 4 s, and was processed at a central location (an on-site machine at University College Dublin). Future versions of SBCI-Watch will likely feature edge computing architecture, after the privacy implications of such architecture on the use case has been ascertained.

After data pre-processing, we realized over 2000 captures. Based on the information which we gathered from the literature [7] and based on the observed ground truth, we capped the chaotic runs at 15 km/h while walking/stationary under calm situations were capped at about 7 km/h; outliers were removed based on this. The data was also labelled and reshaped accordingly.

Choice of Deep Learning and Hierarchical HAR Approach: Traditional Machine Learning methods are generally being overtaken by deep learning methods in many areas. The success of Recurrent Neural Networks (RNNs), and the LSTM in particular, in handling input data that has temporal patterns is well documented in the literature [28]. Accordingly, LSTM has been used for our model; the loss generally drops after relatively few iterations and the predictions are in line with the ground truth which was observed in the field by the researcher. LSTM ensures that data is not simply classified, it helps our model to account for circumstances within the area being monitored, prior to the time of monitoring; without such consideration, the analysis will not make much sense.

Furthermore, we adopted the hierarchical (HAR) approach which has also been reported to have improved performance [34]. The main idea behind this concept is to have heterogeneous and hierarchical layers or classifiers that deal with specific use-cases using more relevant or impactful features, and then similarly moving on to applicable sub-categories, leading to a more precise result. The downside of this approach is a possibility of higher computational load, but it has also been shown that this is not a big issue in practice [28]. And in the case of LSTM, the number of layers for each sub-category can be carefully chosen to better conserve resources.

As illustrated in Fig. 5, Our model first uses generic information such as movement pattern, time, and previous state (location is also a factor here) to determine whether there is a case of public disturbance in a public place, for instance a museum or a given location in a smart city street. If there is a significant reason to suspect that a serious act of public disturbance has occurred or is currently taking place, then a further verification step is taken in a sublayer of the hierarchy, combining the earlier output with additional more specific data, which in this case is the feature that was extracted from a shooting detection dataset, published by Khan et al. [33]. This combination is fed back into the model and further processed for a better precision. Basically, the sub-layer seeks to know if there are specific serious crimes that have also been detected in the area at about the same time. This may help law enforcement, for instance, to decide whether an action is necessary as well as the kind of response that may be needed. Such details may not be available with a traditional nonhierarchical approach. To calculate the resultant energy (E) from the x, y, and z axis of the accelerometer data, we used the same formula that was used in [33]:

$E = \sum_{i=0}^{l} = (x^2 + y^2 + z^2)$, where l represents the window length. We used l = 5, to correspond to the 20 s time steps which the model uses. Recall that our dataset was recorded per 4 s, during the data collection. Khan et al. [33] uses this energy to select gunshot related frames, we refer readers to their work for more on this.

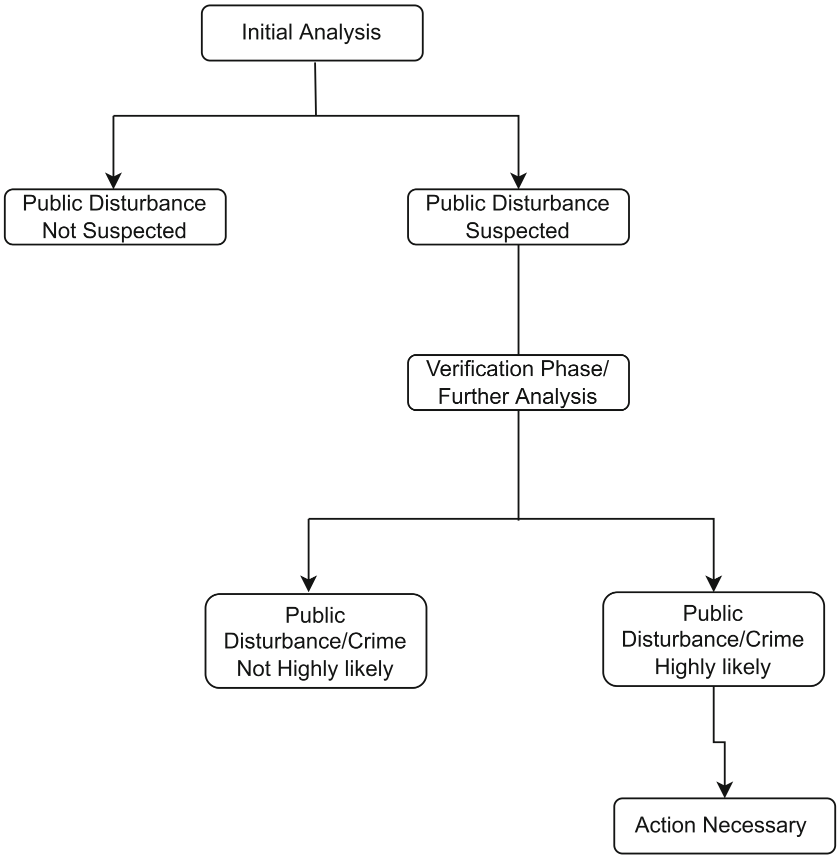

Fig. 5. The Hierarchical HAR Approach used in SBCI-Watch

Results and Discussion: The data set was split for training (90%) and testing (10%). The split was done sequentially because using the traditional train_test_split to do it (randomly) would be unsuitable. We need to keep note of prior events, and splitting the data randomly would make it disorderly. LSTM was used because of its documented success with this kind of data.

The model converges smoothly with an accuracy of about 98.5%. Accuracy ranges between 97.3 and 99.8% were also recorded. Our result is generally better than the results of most HAR-related reports which may be considered to be in the same domain as our work, including 95.21% which was achieved by Yang et al. [31], 98% by zhou et al. [32], and 97.2% recorded by Jackermeie and Ludwig [28]. Jackermeie and Ludwig also recorded performances that ranged from 62.6% to 98.7%, but our model is still within their best recorded performance.

Our goal is to be able to identify when people start 'running for their lives' at a location or venue within a smart city, e.g. Museum. Figure 6 illustrates how the model performed in the task. People walk/run at different paces and for different reasons, there were also other factors considered, such as the timing and location, among others. Chaos or public disturbance may be suspected if there is a clear deviation from the normal state of pedestrian traffic in an area, and people appear to be running in large numbers in different directions and seemingly deserting (or fleeing from) an area, as shown in Fig. 3. The deserted area(s) could be the points of attack, and with close to real-time information processing, the law enforcement may be able to respond swifter, for instance, by deploying drones and personnel quickly to the scene.

In general, our model was able to reconcile the differences; identifying ideal and panic situations with a good level of accuracy, in line with the ground truth. The data was reshaped accordingly, and the model steps back every 20 s. This enables it to make predictions regarding the likely current situation of a location. There is no particular reason for choosing the 20 s interval, it is merely for illustration purposes. The best fit (interval) may be subject to the peculiarities of a location being monitored and available resources. That is, a threshold that may be considered high enough to trigger an alarm can be set at the decision-making level, considering possible peculiar security needs of a given area or location.

The unexpected spike which can be noticed in Fig. 7 may be due to the fact that the model does not have an earlier reference point, since it was still at an 'initialization' stage when the spike occurred. Nonetheless, decision making based on this model, particularly with regard to the considered scenario, would be relatively accurate because it was able to predict the occurrences in line with the ground truth. It is however important to understand that more data regarding a given venue (considering different scenarios) may need to be gathered over a period of time, and fed into the model, in order to determine the day-to-day situation of that venue, and to allow for higher confidence in the outcome, particularly in a production environment. This work is therefore a proof of concept at this stage.

The presented results capture several sessions which were used in the testing phase. As earlier explained, during each session, all the participants perform two main acts; walking normally (or/and other normal activities including staying idle), and running as if they were in panic. Data were then gathered and labelled according to the ground truth. Figure 7 illustrates the sessions.

Figure 8 represents the result of the additional verification step which the hierarchical approach has allowed us to do. Results from the previous step were combined with the firearm data which were earlier described, in order to ascer-

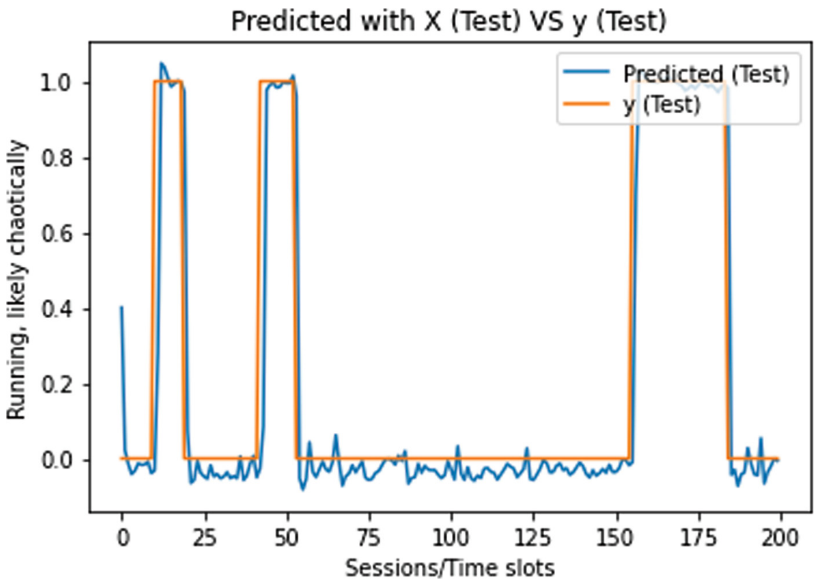

Fig. 6. Sessions of simulated panic running. People walk and run at different paces among other features, but the model is able to identify the panic sessions.

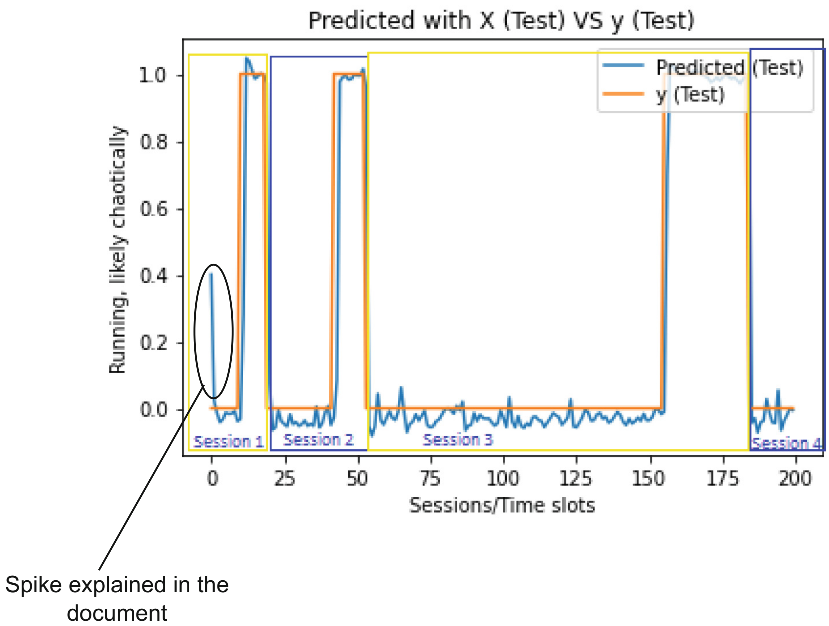

Fig. 7. Showing different sessions, and a slight anomaly

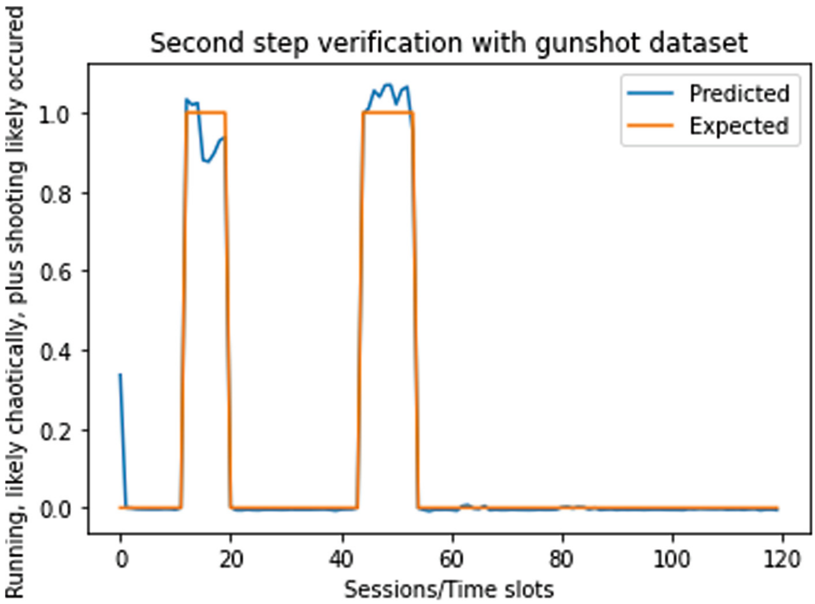

Fig. 8. Further verification step, using firearm data-set. This kind of verification will help to minimise false alarms.

tain if there was also an indication of shooting, in addition to the suspected scene of public disturbance. This would add to the weight and confidence in the recommendation being made. Like in the previous results, the clearly high spikes or peaks indicate cases of suspected public disturbance, incident, or a related crime at the location being considered. The clearly low ones represent normal activity, while the not-so-high spike at the beginning, which was also recorded in the previous results, can be ignored for the same reason which was earlier stated. We would like to state that the two datasets (the one generated by us, and the one from literature) were obviously not collected at the same time and location, therefore such information (about time and location) was abstracted out, particularly in the second step.

4 Conclusion and Future Work

Various aspects of our cities and homes are becoming smarter. In some cities, drones and some kinds of robots are already able to deliver packages to the right addresses. We can warm up our homes even without being physically there, and many people are now familiar with the idea of autonomous cars. But this is just a tip of the iceberg. Different components of the cities are expected to continue to evolve along this line, including crime watch, policing, and response to distress.

This work has explored the prospects of a sensor-based surveillance system for a smart city, focusing on a particular location. We have demonstrated that common sensors which are available on devices that we use daily, such as smartphones and smartwatches can be used to gain insight into the movement patterns of individuals in a location or venue such as a Museum. And that based on such information and other relevant data, which can also be gathered through the sensors (such as gunshot information), it is possible to predict incidents of public disturbances or crime automatically, so that appropriate authorities can be likewise alerted for swift intervention. This sensor-based approach can be an alternative, or a compliment, to the vision-based approach which can be more expensive and more privacy intrusive.

We have focused on using readily available Apps and devices of various types to make adoption easy for any entity that may consider adopting SBCI-Watch, when it is fully developed. The data which may be collected from consenting individuals is almost or exactly the same data that they are already willing to share while using their regular Apps. The custodian of such data is expected to remain the same, to ensure necessary compliance. For example, many Museums already have Apps which they use - if they were to adopt SBCI-Watch, it will not take control or responsibility over the data away from them; they would still be in-charge.

Results suggest that the solution is viable and we believe that it can help to tackle terrorism and other serious crimes in the public spaces (of a smart city). It would help appropriate authorities to get information and respond swifter to incidents within the public spaces. It is also cost-effective, since most of the resources needed to make it work are embedded into the city.

Future Work: We shall be expanding the scope of the project, having ascertained its viability through this work. We shall seek to collaborate with the managers and other individuals within Museums and similar public spaces around Ireland and Europe to bring about a bigger scope. This shall be done with full compliance to applicable regulations related to data and a keen adherence to data privacy good practices.

Acknowledgement. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 883596. The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

References

1. Dhiman, C., Vishwakarma, D.K.: A review of state-of-the-art techniques for abnormal human activity recognition. Eng. Appl. Artif. Intell. **1**(77), 21–45 (2019)
2. Fereidoonian, F., Firouzi, F., Farahani, B.: Human activity recognition: from sensors to applications. In: 2020 International Conference on Omni-layer Intelligent Systems (COINS), pp. 1–8. IEEE (2020)
3. Zhang, L., Cui, W., Li, B., Chen, Z., Wu, M., Gee, T.S.: Privacy-preserving cross-environment human activity recognition. IEEE Trans. Cybern. (2021)

4. Song, B., Sheng, R.: Crowd counting and abnormal behavior detection via multi-scale GAN network combined with deep optical flow. Math. Probl. Eng. **16**, 2020 (2020)
5. Pang, L.X., Chawla, S., Liu, W., Zheng, Y.: On detection of emerging anomalous traffic patterns using GPS data. Data Knowl. Eng. **1**(87), 357–373 (2013)
6. Pan, B., Zheng, Y., Wilkie, D., Shahabi, C.: Crowd sensing of traffic anomalies based on human mobility and social media. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 344–353 (2013)
7. Faye, S., Frank, R., Engel, T.: Adaptive activity and context recognition using multimodal sensors in smart devices. In: Sigg, S., Nurmi, P., Salim, F. (eds.) MobiCASE 2015. LNICST, vol. 162, pp. 33–50. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-29003-4_3
8. Irfan, M., Tokarchuk, L., Marcenaro, L., Regazzoni, C.: Anomaly detection in crowds using multi sensory information. In: 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–6. IEEE (2018)
9. Hoogendoorn, S.P., Bovy, P.H.: Pedestrian route-choice and activity scheduling theory and models. Transp. Res. Part B: Methodol. **38**(2), 169–190 (2004)
10. Fruin, J.J.: Designing for pedestrians a level of service concept. Polytechnic University (1970)
11. Yang, L., Rao, P., Zhu, K., Liu, S., Zhan, X.: Observation study of pedestrian flow on staircases with different dimensions under normal and emergency conditions. Saf. Sci. **50**(5), 1173–1179 (2012)
12. Wang, B., Ye, M., Li, X., Zhao, F., Ding, J.: Abnormal crowd behavior detection using high-frequency and spatio-temporal features. Mach. Vis. Appl. **23**(3), 501–511 (2012)
13. Favaretto, R.M., Dihl, L.L., Musse, S.R.: Detecting crowd features in video sequences. In: 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 201–208. IEEE (2016)
14. Duives, D.C.: Analysis and modelling of pedestrian movement dynamics at large-scale events. Doctoral dissertation, Delft University of Technology (2016)
15. Centorrino, P., Corbetta, A., Cristiani, E., Onofri, E.: Measurement and analysis of visitors' trajectories in crowded museums. arXiv preprint arXiv:1912.02744 (2019)
16. Versichele, M., Neutens, T., Delafontaine, M., Van de Weghe, N.: The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: a case study of the Ghent Festivities. Appl. Geogr. **32**(2), 208–220 (2012)
17. Danalet, A., Tinguely, L., de Lapparent, M., Bierlaire, M.: Location choice with longitudinal WiFi data. J. Choice Modell. **1**(18), 1–7 (2016)
18. Gioia, C., Sermi, F., Tarchi, D., Vespe, M.: On cleaning strategies for WiFi positioning to monitor dynamic crowds. Appl. Geom. **11**(4), 381–399 (2019)
19. Ton, D., van den Heuvel, J., Daamen, W., Hoogendoorn, S.: Route and activity location choice behaviour of departing passengers in train stations. In: hEART (European Association for Research in Transportation) 2015 Conference. Copenhagen, Denmark, pp. 9–11 (2015)
20. van der Spek, S.: Spatial metro: tracking pedestrians in historic city centres. Res. Urban. Ser. **1**(1), 77–97 (2008)
21. Blanke, U., Tröster, G., Franke, T., Lukowicz, P.: Capturing crowd dynamics at large scale events using participatory GPS-localization. In: 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pp. 1–7. IEEE (2014)

22. Duives, D.C., Wang, G., Kim, J.: Forecasting pedestrian movements using recurrent neural networks: an application of crowd monitoring data. Sensors **19**(2), 382 (2019)
23. Zhang, K., Wang, M., Wei, B., Sun, D.: Identification and prediction of large pedestrian flow in urban areas based on a hybrid detection approach. Sustainability **9**(1), 36 (2016)
24. Yang, Y., Heppenstall, A., Turner, A., Comber, A.: Who, where, why and when? Using smart card and social media data to understand urban mobility. ISPRS Int. J. Geo Inf. **8**(6), 271 (2019)
25. Hoseinzadeh, N., Arvin, R., Khattak, A.J., Han, L.D.: Integrating safety and mobility for pathfinding using big data generated by connected vehicles. J. Intell. Transp. Syst. **24**(4), 404–420 (2020)
26. Chen, C., Zhang, D., Samuel Castro, P., Li, N., Sun, L., Li, S.: Real-time detection of anomalous taxi trajectories from GPS traces. In: Puiatti, A., Gu, T. (eds.) MobiQuitous 2011. LNICST, vol. 104, pp. 63–74. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30973-1_6
27. Jung, I.Y.: A review of privacy-preserving human and human activity recognition. Int. J. Smart Sens. Intell. Syst. **13**(1), 1–3 (2020)
28. Jackermeier, R., Ludwig, B.: Smartphone-based activity recognition in a pedestrian navigation context. Sensors **21**(9), 3243 (2021)
29. Joh, E.E.: Policing the smart city. Int. J. Law Context **15**(2), 177–182 (2019)
30. Welsh, D., Roy, N.: Smartphone-based mobile gunshot detection. In: 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 244–249. IEEE (2017)
31. Yang, J., Cheng, K., Chen, J., Zhou, B., Li, Q.: Smartphones based online activity recognition for indoor localization using deep convolutional neural network. In: 2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS), pp. 1–7. IEEE (2018)
32. Zhou, B., Yang, J., Li, Q.: Smartphone-based activity recognition for indoor localization using a convolutional neural network. Sensors **19**(3), 621 (2019)
33. Khan, M.A., Welsh, D., Roy, N.: Firearm detection using wrist worn tri-axis accelerometer signals. In: 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 221–226. IEEE (2018)
34. Ashqar, H.I., Almannaa, M.H., Elhenawy, M., Rakha, H.A., House, L.: Smartphone transportation mode recognition using a hierarchical machine learning classifier and pooled features from time and frequency domains. IEEE Trans. Intell. Transp. Syst. **20**(1), 244–252 (2018)